MethodsX 9 (2022) 101717

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

From single to multivariable exposure models to translate climatic and air pollution effects into mortality risk. A customized application to the city of Rome, Italy ☆,☆☆

M. Michetti [a,*], M. Adani [a], A. Anav [b], B. Benassi [c], C. Dalmastri [c], I. D'Elia [d], M. Gualtieri [a], A. Piersanti [a], G. Sannino [b], R. Uccelli [c], G. Zanini [a]

[a] *Division of Models and Technology for Risk Reduction, ENEA Centro Ricerche Bologna, Via Martiri di Monte Sole 4, Bologna 40129, Italy*

[b] *Division of Models and Technology for Risk Reduction, ENEA Centro Ricerche Roma Casaccia, Via Anguillarese 301, Rome, Santa Maria di Galeria 00123, Italy*

[c] *Division of Health Protection Technologies, ENEA Centro Ricerche Roma Casaccia, Via Anguillarese 301, Rome, Santa Maria di Galeria 00123, Italy*

[d] *Division of Models and Technology for Risk Reduction, ENEA Centro Ricerche Roma, Lungotevere Thaon de Revel, 76, Rome 00196, Italy*

A B S T R A C T

This study presents an approach developed to derive a Delayed-Multivariate Exposure-Response Model (D-MERF) useful to assess the short-term influence of temperature on mortality, accounting also for the effect of air pollution (O_3 and PM_{10}). By using Distributed, lag non-linear models (DLNM) we explain how city-specific exposure-response functions are derived for the municipality of Rome, which is taken as an example. The steps illustrated can be replicated to other cities while the statistical model presented here can be further extended to other exposure variables. We derive the mortality relative-risk (RR) curve averaged over the period 2004–2015, which accounts for city-specific climate and pollution conditions.

Key aspects of customization are as follows:

This study reports the steps followed to derive a combined, multivariate exposure-response model aimed at translating climatic and air pollution effects into mortality risk.

☆ Direct Submission or Co-Submission: Co-Submission

☆☆ STOTEN-D-21-28734

DOI of original article: 10.1016/j.scitotenv.2022.1546803

* Corresponding author.

E-mail address: melania.michetti@enea.it (M. Michetti).

https://doi.org/10.1016/j.mex.2022.101717

Integration of climate and air pollution parameters to derive RR values.

A specific interest is devoted to the investigation of delayed effects on mortality in the presence of different exposure factors.

ARTICLE INFO
Method name: Delayed-Multivariate Exposure-Response Functions: D-MERF
Keywords: DLNM, Relative Risk, Temperature, Air pollution, Time-pattern analysis, Delayed effect investigation, Integrated exposure model, D-MERF
Article history: Received 6 December 2021; Accepted 22 April 2022; Available online 5 May 2022

Specifications table

Subject Area:	Environmental Science
More specific subject area:	*Health assessment*
Method name:	*Delayed-Multivariate Exposure-Response Functions: D-MERF*
Name and reference of original method:	***DLNM function description:*** *Gasparrini A., Armstrong B., Kenward M.G., 2010. Distributed lag non-linear models. Stat Med. 2010 Sep 20; 29(21):2224-34. DOI: 10.1002/sim.3940. PMID: 20812303; PMCID: PMC2998707.* ***Code for models comparison*** *in the supplementary material of Armstrong B., Sera F., Vicedo-Cabrera A.M., Abrutzky R., OudinÅström D. et al., 2019. The Role of Humidity in Associations of High Temperature with Mortality: A Multicountry, Multicity Study. Environ. Health Persp.127 (9) doi: 10.1289/EHP5430.* ***Tutorial for RR derivation****: Vicedo-Cabrera AM, Sera F, Gasparrini A. Hands-on Tutorial on a Modeling Framework for Projections of Climate Change Impacts on Health. Epidemiology. 2019 May; 30(3):321-329. DOI: 10.1097/EDE.0000000000000982. PMID: 30829832; PMCID: PMC6533172.* ***Time series investigation:*** *Bhaskaran K., A. Gasparrini, S. Hajat, L. Smeeth and B. Armstrong, 2013. Time series regression studies in environmental epidemiology. International Journal of Epidemiology 2013; 42:1187–1195 DOI:10.1093/ije/dyt092*
Resource availability	**Sources for raw databases retrieval:** *Pollution data:* https://www.eea.europa.eu/data-and-maps/data/aqereporting-9 *Temperature data:* https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form *Mortality data:* https://www.istat.it/it/archivio/240401 ***Specific time series used:*** *Temperature, pollution, and mortality time series elaborated in this article are available in the Supplementary Material* ***Model Code:*** *the modeling chain used in the context of this exercise can be eventually made available upon request to the corresponding author.* ***Major R libraries used to run the analysis:*** *dlnm*: https://cran.r-project.org/web/packages/dlnm/index.html *splines:* https://rdocumentation.org/packages/splines/versions/3.6.2 *jtools*: https://rdocumentation.org/packages/jtools/versions/2.2.0 *ggplot2:* https://rdocumentation.org/packages/ggplot2/versions/3.3.5 *dplyr:* https://rdocumentation.org/packages/dplyr/versions/0.7.8 *stats*: https://rdocumentation.org/packages/stats/versions/3.6.2 *lubridate:* https://rdocumentation.org/packages/lubridate/versions/1.8.0 *metan:* https://rdocumentation.org/packages/metan/versions/1.16.0

Data description

We constructed a database for the city of Rome referring to the period 2004 and 2015, combining all-natural causes of mortality with climate and air pollution data. In statistical terms, mortality data represent our response variable while climatic and air pollution information characterize our regressors (independent variables) in the model equations below described. More specifically:

- Daily number of deaths can be requested from the Italian National Institute of Statistics (ISTAT), which is the responsible agency for the codification of death certificates in Italy. Due to privacy

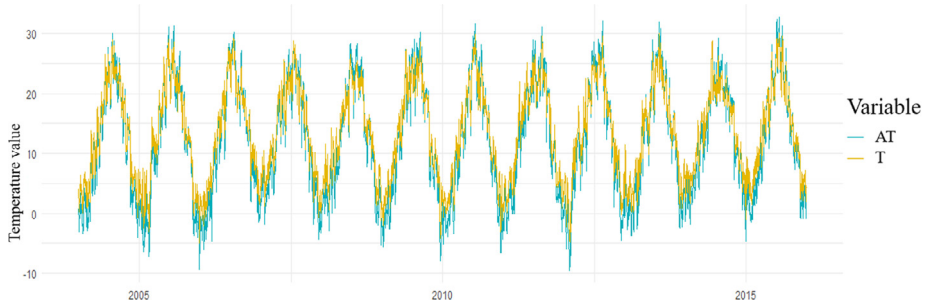

Fig. 1. Mean daily temperature (T) and apparent mean daily temperature (AT) from 2004 to 2015 (°C).

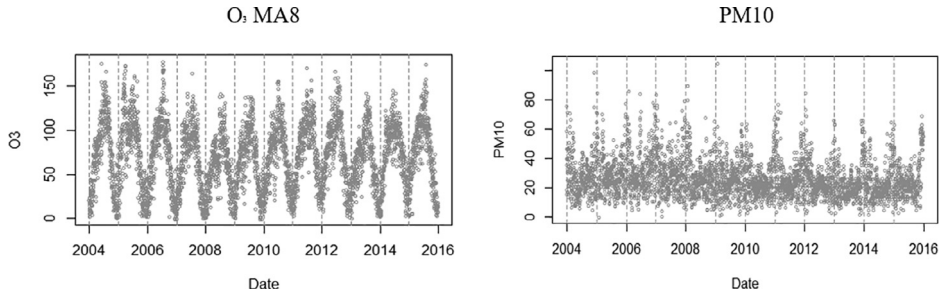

Fig. 2. O_3 MA8 and PM_{10} across 2004–2015 ($\mu g/m^3$).

issues, such data are provided by ISTAT to ENEA in anonymous form for epidemiological evaluation purposes, in the framework of the Italian National Statistical System (SISTAN) network.

- As for the temperature exposure factor, we used two indicators for robustness check and result validation: the daily mean temperature (T) and the apparent temperature indicator (AT); the latter approximates the heat balance in the human body by combining humidity, wind speed and temperature in a linear relation [5]. T data result from the elaboration of hourly surface air temperatures, retrieved from the European Centre for Medium-range Weather Forecast ERA5 reanalysis [16], having a horizontal resolution of ∼31 km from 1979 onwards. Similarly, AT values are calculated from the ERA5 database, following Steadman [26] and Buzan et al. [8] criteria (Fig. 1).

Particulate matter (PM_{10}) and Ozone (O_3) time series for the IT0953A background station (Roma - Villa Ada) are retrieved from the European database [1]. The daily maximum 8 h moving average for O_3 (MDA8) and the daily mean for PM_{10} are calculated as daily parameters relevant for the exposure - response functions calculations (O_3 MA8 and PM_{10} in Fig. 2).

Method details

Approach overview

We present the approach followed to explore the combined short-term impact of mean temperature (T) and apparent mean temperature (AT), and air pollution (PM_{10} and O_3) on mortality from all-natural causes, using daily time-series for the Italian municipality of Rome.

Distributed lag nonlinear models (DLNMs, [13]) are used to generate the basis for our Delayed-Multivariate Exposure-Response Function (D-MERF). The DLNM approach is widely recognized to evaluate lagged effects of several environmental factors on a specific health outcome and to reveal complex association between exposure and response variables [17].

In the presence of daily count data for an outcome variable (i.e. all natural causes mortality in our study), related information can be assumed to originate from a Poisson distribution, which is overdispersed, and a log-link with the selected outcome. More in detail, Eq. (1) describes our generalized Poisson model used for the estimation of the population exposure to the combined effect of temperature and air pollution. The convenience of using a Poisson regression (with overdispersion) model for count data lies in the opportunity of interpreting the regression coefficients of the explanatory variables as the associated relative risk (RR). In Eq. (1), Y_t is the mortality count in day t, $E(Y_t)$ is the corresponding expectation of the Poisson distribution, α is the intercept and β is the vector of the regression coefficients associated to each covariate $Z_{t,l}$. The covariates here represent the values of mean temperature (either T or AT) and the concentration of air pollutants (both O_3 and PM_{10}) at lag day (l) with l $(0, L)$ that can have different ranges over different study sites (cities, regions, etc.).

$$Log[E(Y_t)] = \alpha + \beta Z_{t,l} + ns(time, df) + \gamma DOW_t + \varkappa Month_t \quad (1)$$

The best model design according to Eq. (1) should bring together the most efficient exposure-response and lagged specifications derived by employing single-variable models for each exposure variable. To this aim, based on Eq. (1), we investigate several functional forms and derive a specific distributed lag model with different lag structures and effects for each of the exposure-response relationships (T/AT-mortality, O_3-mortality, and PM_{10}-mortality). Additionally, we account for time trends and cyclicality, seasonality patterns, and irregular effects. Specifically, to capture the effect of time-varying confounders we include a natural spline smoothing function of time - the *ns(time, df)* - and model the categorical variables representing the day of the week (Dow_t) and the month of the year ($Month_t$). We also test the lag responses within a 30-day period (one month), for each exposure variable.

Below, we present the steps leading to the definition of the D-MERF, which combines different exposure factors and accounts for different delayed effects on the outcome variable.

Step One. Time patterns analysis

Our analysis is aimed at investigating short-term effects of temperature and air pollution on mortality, i.e., whether short-term changes in our outcome variable (death number) can be explained by short-term variations in the exposure factors (daily temperatures and air pollution levels). A key aspect is represented by potential autocorrelation problems between observations and long-term patterns that may confound our short-term evaluation. Based on the relations pictured in **Fig. 3** we have evidence of autocorrelation among the observations characterizing the outcome variable.

The existence of autocorrelation requires testing for several time patterns, such as long-term trend, cycle and seasonality, monthly and day-of-week patterns, and holiday/calendar effects. When different exposure variables are in place, the testing should be repeated for each one: the daily maximum 8h running mean of O_3, the daily mean for PM_{10} and two temperature parameters (T, AT). The removal of potential effects associated with medium-to-long term trends over is achieved by controlling for these effects in the time series, i.e., by testing and including some functions of time in the main model. Once time patterns are controlled for, we are left with short-term effects, unless other confounding factors apply [7].

As for the long-term trend, cycles, and seasonality, we fit different functions of time: time-stratified model, periodic functions and spline functions formed of polynomial segments. The shape of the predicted deaths, reported in red in **Fig. 4,** highlights the inability of time-stratified models and periodic functions (Fourier terms) to reproduce properly the variability of mortality across time (plots **A** and **B**).

Among spline functions of time, we test a cubic B-spline (piecewise third-order polynomials that pass through a set of control points) and natural cubic splines (which extrapolates linearly beyond the boundary knots). Results achieved using spline functions outperform those associated to time-stratified and periodic models (plots **C to F**). Despite predicted mortality between the cubic B-spline (plot **C**) and the natural-spline functions (plots **D to F**) appear comparable, our choice falls on the use of natural splines, whose structure is based on piecewise interpolation of low-degree polynomials

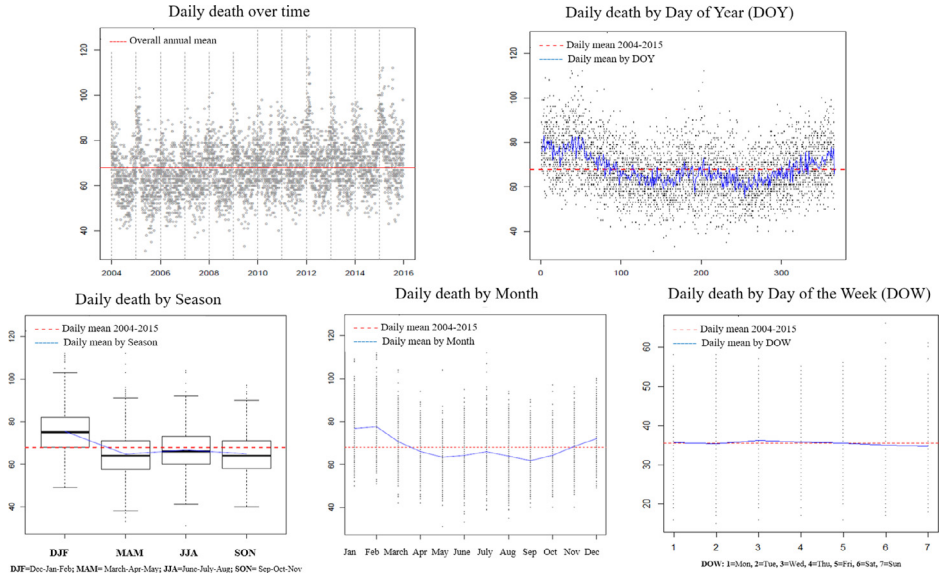

Fig. 3. Medium to-long term patterns of the outcome variable (Rome case study)
Notes: The y-axes reports the number of daily death in all graphs.

Table 1
Variation (%) in estimation parameters across different natural spline functions.

Time intervals	Models compared	DF	Dispersion	Deviance	AIC	Pseudo-R^2
Mortality expressed as a function of mean Apparent Temperature						
8 vs 6	98 vs 74 DF	32.43	- 1.91	- 2.43	- 1.36	4.17
12 vs 8	146 vs 98 DF	48.98	- 2.26	- 3.18	- 1.05	2.00
Mortality expressed as a function of mean Temperature						
8 vs 6	98 vs 74 DF	32.43	- 1.97	- 2.49	- 1.42	2.08
12 vs 8	146 vs 98 DF	48.98	- 2.20	- 3.13	- 1.00	4.08

Notes: we compare the percentage variation in DF, Dispersion, Deviance, AIC and Pseudo-R^2 of models with different time interval specifications, reported under the column "Models compared".

for each time interval. Compared to cubic B-splines, this allows using a lower number of degrees of freedom (number of knots), thereby limiting oscillations and non-convergence issues, compared to higher order-degree polynomials.

To infer the correct level of flexibility of the spline function and since there is no consensus on the number of optimal knots, we develop a simple sensitivity analysis on different assumptions on time-intervals, by testing natural splines with 6, 8, 12-time intervals per year, using respectively 74, 98, and 146 degrees of freedom (DF), including boundary knots. At this stage, mortality is analyzed as a function of AT/T only. The results on the sensitivity are presented for models including only the temperature variable, while air pollution factors are added as additional exposure parameters in a second stage. **Table 1** shows the percentage variation in several estimation parameters as we pass from a lower to a progressively higher flexibility of the spline function that requires the estimation of additional coefficients, i.e., a greater investment in degrees of freedom.

The choice of the best model should aim at achieving a proper representation of seasonality and time trends while avoiding over-specification problems and leaving enough information to estimate the exposure effects [7]. While dividing the year into 12 periods would contribute to reducing the dispersion, deviance and the Akaike Information Criterion (AIC), these reductions are counterbalanced

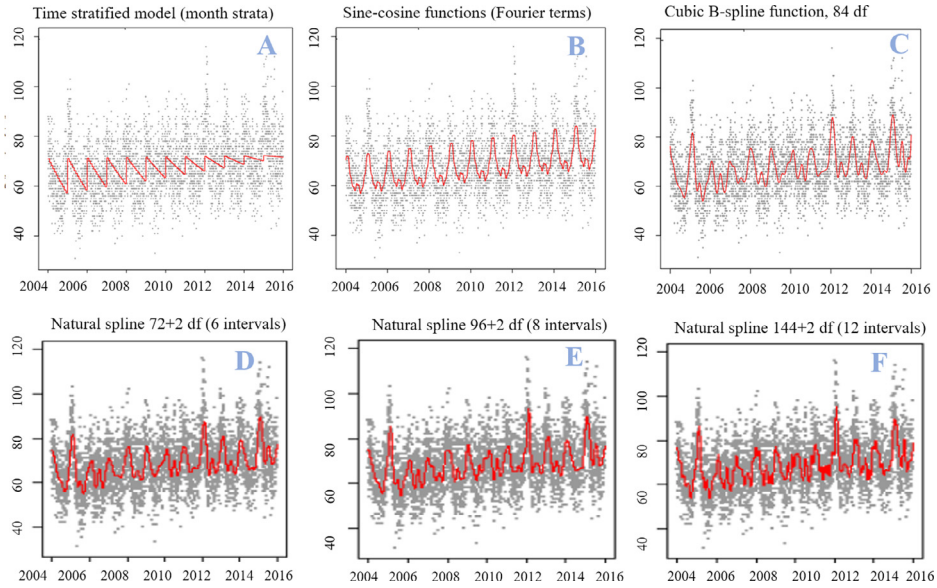

Fig. 4. Different approximation functions to model long-time trends for the response variable (predicted number of deaths in red).

by an increase in the degrees of freedom used and of a very contained effect in terms of higher pseudo-R^2. We judge therefore 8-time intervals to be sufficient to represent mortality variability.

We also test for holidays effects (dummy variable assuming value of '1' for public holidays in Rome) and seasonality expressed in different ways (climatological seasons: June, July, August -JJA; September, October, November – SON; December, January, February – DJF; March, April, May - MAM; summer and winter seasons) but none of the seasonality representation show a statistically significant time pattern. The case of weekly trends is different, as it shows a significant effect of specific days of the week (DOW) on mortality. Indeed, health risk could be influenced by industrial, economic and social activities, including weekly people's behavior that may affect pollution concentrations. We finally add a control for months (*Month*: categorical variable ranging 1–12) to account for monthly patterns.

Table 2 compares results across different model specifications (either accounting for - or neglecting - time patterns) and shows the related goodness of fit indicators; the best models are highlighted in bold. Here, mortality is again analyzed as a function of AT/T, making assumptions neither on the functional relation between temperature and health outcome, nor on the existence of delayed effects. The resulting best model equations, highlighted in bold in **Table 2** include a natural spline function of time, the month pattern and the DOW variables, which are statistically significant and generate a substantial improvement in predicted mortality (see **Eq. (1)**).

Step two. Explanatory variables and their relative role

Statistically, in short-term analyses temperature seems to exert a greater effect on daily deaths when compared to air pollutants as it can worsen the mortality picture not only in summer but also in winter. To assess the relative contribution of pollution variables we compare four models. As for section 1.1, we are not yet making assumptions on the functional relations between exposure variables and mortality and on their delayed effects on mortality.

Among the four models investigated, the first accounts only for temperature (either daily mean or daily apparent mean), the second and the third include also O_3 and PM_{10}, respectively, while the

Table 2
Goodness of fit parameters and estimation coefficients (RR) for adjusted and unadjusted models for both temperature variables (AT and T).

		Daily mean apparent temperature-AT							
N	Model	Obs	DF	Dispersion	Deviance	AIC	Coeff.	ci.low	ci.high
1	Unadjusted	4383	2	1.67	7286	7293	-0.571	-0.625	-0.517
2	Long_term	4383	98	1.17	4994	5222	0.857	0.719	0.995
3	Month	4383	99	1.16	4989	5219	0.842	0.703	0.981
4	Dow	4383	99	1.16	4967	5196	0.863	0.725	1.001
5	Dow2	4383	104	1.15	4959	5200	0.862	0.724	0.999
6	**Dow2_month**	**4383**	**105**	**1.15**	**4954**	**5197**	**0.846**	**0.708**	**0.985**
		Daily mean temperature-T							
1	Unadjusted	4383	2	1.67	7259	7265	-0.733	-0.801	-0.665
2	Long_term	4383	98	1.16	5007	5236	1.086	0.904	1.269
3	Month	4383	99	1.16	5001	5232	1.067	0.884	1.250
4	Dow	4383	99	1.16	4980	5210	1.094	0.912	1.276
5	Dow2	4383	104	1.15	4972	5214	1.093	0.911	1.275
6	**Dow2_month**	**4383**	**105**	**1.15**	**4966**	**5210**	**1.073**	**0.891**	**1.256**

Notes: RR stands for relative risk estimations that switch from negative to positive values as we account for time-related confounding factors; they represent therefore the magnitude of the estimated coefficients, by model, when the Apparent Daily Mean Temperature (AT) and Daily Mean Temperature (T) are considered, respectively. The ci.low and ci.high are the low and high confidence intervals. **Models' legend:** Unadjusted (not accounting for time trends/other confounding factors); Long_term (including a natural spline function of time); Month (also accounting for monthly variability); Dow & Dow2 (adding different specifications of the day of the week indicator); Dow2_month (final model, in bold because more performant).

Table 3
Risk coefficients of tested models and confidence intervals.

Model	Variable coefficient	Mean Apparent Temperature (AT)			Mean Temperature (T)		
		Estimate	ci.low	ci.high	Estimate	ci.low	ci.high
Only-AT/T	AT/T	0.846***	*0.708*	*0.985*	1.073***	*0.891*	*1.256*
AT/T+O_3	AT/T	0.827***	*0.688*	*0.967*	1.046***	*0.861*	*1.232*
	O_3 MA8	0.026*	*0.003*	*0.049*	0.018	*-0.006*	*0.041*
AT/T+PM_{10}	AT/T	0.779***	*0.633*	*0.924*	0.989***	*0.803*	*1.176*
	PM_{10}	0.054**	*0.018*	*0.089*	0.074***	*0.039*	*0.108*
AT/T+O_3+PM_{10}	AT/T	0.752***	*0.605*	*0.899*	0.952***	*0.762*	*1.142*
	O_3 MA8	0.058**	*0.022*	*0.093*	0.077***	*0.042*	*0.111*
	PM_{10}	0.030*	*0.007*	*0.053*	0.023	*-0.001*	*0.046*

Significance: $p < 0$ '***'; $p < 0.001$ '**'; $p < 0.1$ '*'; $p < 0.05$ '.'; $p < 0.01$ ' '

latter estimates the overall contribution of all exposure variables entering the model simultaneously (**Table 3**).

Adding O_3 to the equation that considers only temperature, slightly lowers the magnitude of the temperature-RR coefficients, which remain however strongly significant ($p = 0.0001$). In addition, only when using the AT parameter the effect of O_3 is significant, and just marginally ($p = 0.01$). This result could be explained, at least partially, by the high correlation between mean daily temperature (and AT) and the daily maximum value of the 8h running means of ozone, calculated over the 24 h ($\rho_{O3\text{-}MA,T} = 70\%$; $\rho_{O3\text{-}MA,AT} = 68\%$). In fact, since warm days are associated with high O_3 levels, the stronger effect of temperature may mask the influence of O_3 on mortality.

Similarly, cold events and increased levels of PM pollution are often associated, so that the effects of one exposure variable may influence the impact of the other. Compared to O_3, PM_{10} is characterized by a higher effect and a higher statistical significance (*p-value* between 0.01 and 0.001) and its inclusion in the temperature-standalone model moves the magnitude of the climatic effect on mortality towards lower values. The overall model conserves the structure of the exposure variables'

Table 4
Goodness of fit of tested models.

N	Model specification	Obs	DF	Dispersion	Deviance	AIC
1	Only AT	4383	105	1.156	4954	5197
2	AT+ O_3MA8	4383	106	1.155	4948	5193
3	AT+PM_{10}	4383	106	1.154	4944	5189
4	**AT+O_3MA8+PM_{10}**	**4383**	**107**	**1.153**	**4937**	**5183**
5	Only T	4383	105	1.159	4966	5211
6	T+ O_3MA8	4383	106	1.158	4964	5210
7	T+PM_{10}	4383	106	1.155	4946	5191
8	**T+ O_3MA8+ PM_{10}**	**4383**	**107**	**1.154**	**4942**	**5189**

relative effects, with temperature remaining the predominant contributor to mortality and PM_{10} and O_3 being not always significant and marginally relevant. In line with the goodness-of-fit results (**Table 4**) the estimation of the RRs relies on the integrated model where all the three exposure variables are considered (highlighted in bold).

Step Three. Lag structure and investigation on the relation between predictors and outcome variables

Tables 3 and **4** refer to models where the lagged dimension of the health effects of temperature and air pollution is not accounted for. However, since both the present and previous days' levels of air pollution/temperature can influence daily mortality values, not considering delayed effects could produce biased estimations. While the existence of delayed exposure effects is a common issue in health risk analysis [15], little evidence exists on the sensitivity of results when different assumptions on the shape and length of the lag structure are in place [15]. Therefore, it is essential bridging this gap using lagged exposure-response associations.

The flexibility of distributed lag nonlinear models [13] used for the present health risk assessment allows including a different specification of the lagged-response dimension for each exposure variable. Hence, we model a specific distributed lag model with different lag structures and effects for each covariate (i.e., the exposure-single models), accounting for time trends and cyclicality, seasonality, holidays, and irregular effects, as previously described.

Single-exposure-variable models – investigating the lagged-exposure association over one month (30 days) – are derived, at first, making no *a priori* assumptions on the exposure-response associations and not adjusting the lagged effects with each other. Hence, we simulate both linear/nonlinear distributional forms according to the specific exposure variable, and unconstrained/constrained lag structures, making use of the concept of "cross-basis matrices" [14].

As for the historical time series of temperature, we can observe a nonlinear pattern (typical inverse J-shaped curve), while for particulate matter we do not have evidence of specific relational shapes. We assume therefore the temperature-mortality relation to be nonlinear and the air pollution-mortality ones to be linear, as also proposed by the extant literature (e.g., [21,22,27,30]. As for constrained lag structures – useful to highlight the existence of autocorrelation in the dataset that makes individual lagged effects confounding each other [7] – these are achieved by lag-stratifying the exposure variables according to graphical evidence and statistical testing. The best lag strata and lag structures are derived by comparing the dispersion parameter, the Pseudo-R^2, the residual deviance, and the Akaike test.

Once the three single-covariate models have been defined, we integrate all the exposure variables in a unique framework, the Delayed-Multivariate Exposure-Response Function (D-MERF), which maintains the single-model assumptions while capturing the variable-specific lagged effects. Indeed, in this final integrated model, we unify the variable-specific lag-stratifications and functional-forms derived.

Single-temperature model

The effects of temperature on health have been claimed to depend on the city location, population age and gender, and on the specific temperature parameter used. When looking at the whole

Table 5
Models tested for the temperature-mortality association.

Model Label		Hypothesis on the relation between temperature and mortality	Lag structure
1	AT/T Model	No *a-priori* hypothesis	Unconstrained
2	UncAT/T Model	Linearity	Unconstrained
3	LagAT/T Model	Linearity	Constrained model (lag-stratified)[1]
4	Ns_AT/T Model	Natural Spline with 3 knots at 10th, 75th, and 90th percentiles of Temperature Distribution	Natural Spline with 3 internal knots.

[1] Implemented stratifications are (2, 13, 14, 15, 16, 17, 24), according to statistical testing performed and graphical visualization.

Table 6
Comparison among models for the temperature-mortality association

N	Model	Obs	DF	Dispersion	Pseudo-R^2	Deviance	AIC
1	AT	4383	104	1.078	0.24	4641	4865
2	UncAT	4353	104	1.089	0.23	4657	4884
3	LagAT	4353	134	1.060	0.26	4500	4784
4	**Ns_AT**	**4353**	**123**	**1.035**	**0.27**	**4400**	**4655**
1	T	4383	104	1.078	0.24	4641	4865
2	UncT	4353	104	1.089	0.23	4659	4886
3	LagT	4353	134	1.061	0.26	4506	4791
4	**Ns_T**	**4353**	**123**	**1.036**	**0.28**	**4405**	**4660**

Notes: See Table 5 for the model labels description

temperature distribution, both cold and heat lagged effects must be captured. This requires testing time-shifted impacts beyond a few days. Specifically, the effects of temperature are estimated to last about 3-4 days for heat (e.g., [10,12,23]) and to persist for longer periods for cold temperatures, up to 30 days [2,4,27]. The best lag structure for temperature-mortality association in Rome is derived by comparing four models for both temperature parameters (AT and T) with different assumptions on the lag structure and exposure-response association (**Table 5**).

Overall, looking at the results and coefficients, all models show that the effect of heat on mortality is stronger than that of cold. As temperature rises, regardless of the chosen model, the effects are more significant and the relative risk increases. As one moves from below '0' to above '0' degrees Celsius, the effects become significant at day '0' and a few days ahead. For cold temperatures, all models confirm longer lasting effects. Results are robust for both parameters despite the magnitude of the lagged effect of T which is higher than that of AT.

According to the results of the tests in **Table 6** and to the Auto and Partial Autocorrelation Functions (plots not reported here), we chose the fourth model (highlighted in bold), which shows significant lagged effects for both climatic parameters lasting up to $\sim$ 15 days, after which we have small and not significant coefficients (**Fig. 5**).

Single-ozone model

The effects of O_3 on health have been claimed to last about 0–1 days [3,27]. However, we test the displacement effect of ozone on mortality for 30 days. We compare four models, as reported in **Table 7**.

Since unconstrained lagged models do not show a clear pattern, we test different stratifications, selected by looking at graphical results of unconstrained models. When using lag-stratified models, results show a marked decline in the ozone effects after the third day (**Fig. 6**).

Independently on the stratification setting, there is evidence of effects at day '0' and few days ahead, after which the effects strongly weaken. In **Table 8**, according to several statistical tests, we show the best lag stratified choice highlighted in bold.

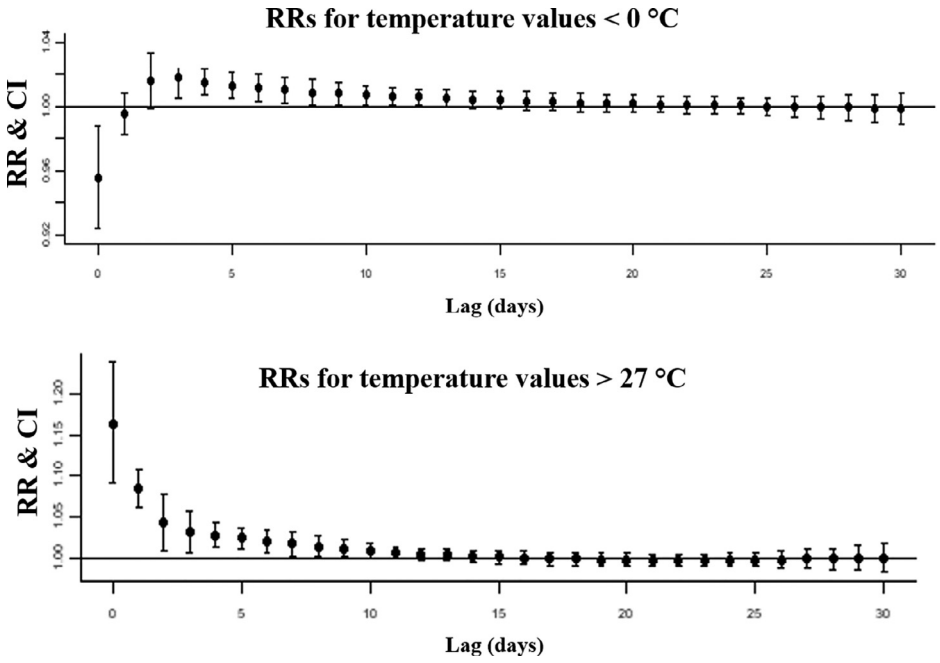

Fig. 5. Lag structures for the temperature-mortality association, for temperature values below '0'°C and above 27°C. Relative risk reported by lag.

Table 7
Models tested for the ozone-mortality association.

Model	Label	Stratifications tested	Hypothesis on the relation ozone - mortality	Lag structure
1	O_3MA8		No *a-priori* hypothesis	Unconstrained
2	Unc_O_3MA8		Linearity + threshold	Unconstrained
3	Strata1_ O_3MA8	4	Linearity + threshold	Constrained model (lag-stratified)
4	Strata2_ O_3MA8	4; 9		

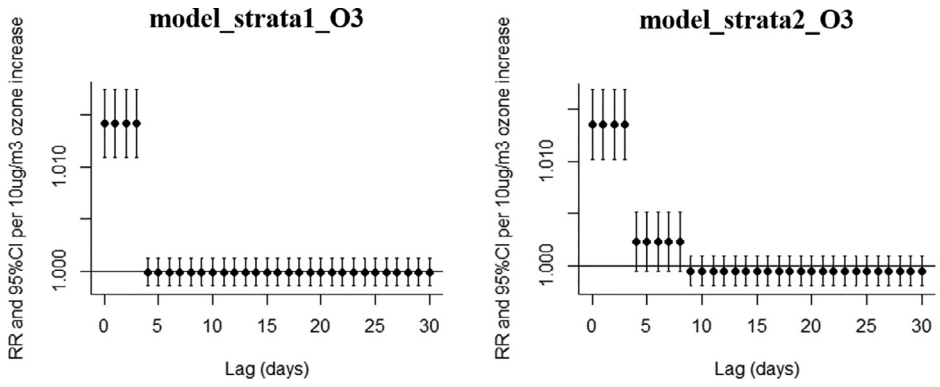

Fig. 6. Comparison among strata models for the ozone-mortality association: RR by lag.

Table 8
Comparison among models for the ozone-mortality association.

N	Model	Obs	DF	Dispersion	Pseudo-R^2	Deviance	AIC
1	O_3MA8	4383	105	1.191	0.36	5107	5357
2	Unc_O_3MA8	4353	130	1.174	0.37	4972	5277
3	Strata1_O_3MA8	4353	104	1.174	0.37	5002	5246
4	**Strata2_O_3MA8**	**4353**	**108**	**1.172**	**0.37**	**4986**	**5240**

Table 9
Models tested for the PM_{10}-mortality association.

Model Label		Stratifications tested	Hypothesis on the relation particulate - mortality	Lag structure
1	PM_{10}		No a-priori hypothesis	Unconstrained
2	Unc_PM_{10}		Linearity	Unconstrained
3	Strata1_PM_{10}	3;9;11;29	Linearity	Constrained model (Lag-stratified)
4	Strata2_PM_{10}	3;10;29		
5	Strata3_PM_{10}	3;10		

Table 10
Comparison among models for the PM_{10}-mortality association.

N	Model	Obs	DF	Dispersion	Pseudo-R2	Deviance	AIC
1	PM_{10}	4353	104	1.196	0.35	5094	5343
2	uncPM_{10}	4353	134	1.175	0.37	4967	5282
3	Strata1_PM_{10}	4353	108	1.177	0.37	5008	5262
4	**Strata2_PM_{10}**	**4353**	**107**	**1.177**	**0.37**	**5008**	**5260**
5	Strata3_PM_{10}	4353	106	1.178	0.37	5013	5263

All the models considered show limited effects of ozone (in magnitude) when seasonality, cycles and other irregular components are accounted for. Interestingly, while we observe positive effects of ozone on excess mortality at lag 0 and at days 1–3, the sign of the effect reverses at longer lags. This outcome could be due to the so-called harvesting effect, which is a common phenomenon within epidemiology, in relation to air pollution (e.g., [25]) and heat waves. The harvesting effect can be conceived as the reduction in mortality occurring immediately after some day(s)/period(s) of excess mortality. The idea behind this phenomenon is that the risk factors (critical temperatures or high levels of pollution) primarily affect people with short life expectancy, i.e., those already affected by one or more diseases, who would have died anyway in a short time. Once susceptible people have died, the remaining stronger/healthier population is less affected and consequently, mortality decreases. In the light of this derivation for the single-ozone model, we include effects at day '0' and within 3-day lags in the overall final model, where other risk factors also apply.

Single-PM10 model

Compared to ozone, less agreement exists on the duration of PM_{10} short-term effects on health and such uncertainty emerges from our results. The lagged effects are reported to extend from 0 to 1 days [2,27], 0–2 days [19], 0–3 days (Stanišić et al. 2016), or even longer periods (0–5 days) according to Scortichini et al [23], depending on age, gender, country, and other factors. Harvesting effects are rarely reported and are found to last up to 10 days (e.g., [24]. As for the case of ozone, we tested different models (**Table 9**) without limiting the number of lagged effects to a few days. Unconstrained models suffer from collinearity and entail a difficult interpretation, even when lagged effects are adjusted with each other. Stratified models, however, just marginally improve statistical tests (**Table 10**) but clarify the lagged structure.

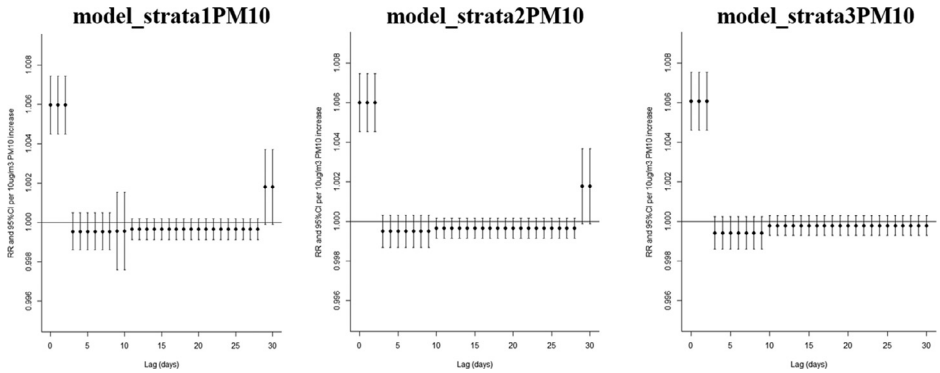

Fig. 7. Comparison among strata models for the PM_{10}-mortality association: RR by lag.

Health effects due to PM_{10} exposure are evident especially at day '0' and up to 2 days beyond. This result applies for all the lag settings when stratification is implemented (**Fig. 7**). Among stratified models, the *Strata2*_PM_{10} represents the best choice. Overall, after controlling for seasonality, holidays' effects, and other time patterns, the PM_{10} effects on mortality are slightly higher than those of O_3.

Step Four. Unifying the single-exposure functions into the D-MERF model

The time pattern and confounding analysis, the displayed effect investigation, and the testing of the exposure-response functional-forms converge into the definition of the D-MERF that combines all the outcomes obtained in previous sections into a final integrated model (**Eq. (2)**). In Eq. (2), the mortality effects (Log[E(**Y**)]) during an average year can be estimated by considering each selected stressors with its multiplying risk coefficient (risk factor).

$$Log[E(Y_t)] = \alpha + \beta_1 Temperature_{t,15} + \beta_2 Ozone_{t,3} + \beta_3 PM10_{t,2} + ns(time, 8) + \gamma DOW_t + \varkappa Month_t \quad (2)$$

Summarizing, following the evidence collected from our elaboration, we modeled the temperature exposure-response curve and the lag-response dimension with a natural spline with three internal knots (placed at the 10th, 75th, and 90th percentiles of the observed temperature distribution for the exposure-response relationship and equally spaced on the log scale for the lag-response one). Delayed effects for temperature were supposed to last up to the 15th day (maximum lag value).

The ozone-mortality association was modeled as a threshold-linear function. To better capture its influence, exerted majorly during summer months (e.g., [31]), the ozone variable was represented as the SOMO35 ppb, i.e., the sum of means over 35 ppb [11]. This corresponds to the inclusion of a threshold of effects for human protection at 70 $\mu g/m^3$ - in line with the interim target of the most recent air pollution guidelines (WHO, [28]) - that allows accounting for the ozone effects during the months April to September, based on our variable distribution. Despite using a linear exposure-response function, describing a process by which the magnitude of the response variable changes as the triggering stimulus exceeds the critical value of 70 $\mu g/m^3$, it is a way, to a certain extent, to model a specific type of nonlinearity, albeit restrictive. Indeed, thresholds act as "knots" and can be described by piece-wise linear segments individually defined by thresholds. Lag effects for ozone were supposed to occur at day '0' and up to the 3rd day.

For PM_{10} and mortality, still modeled with a linear function, we used a conservative approach assuming no lower effect threshold. This is in line with recent literature (e.g., [9]) and the more stringent latest WHO [28] guidelines. In this case, we assumed that the effects of PM_{10} on mortality occur at day '0' and last up to 2 consecutive days.

The outcome of Eq. (2) is a curve of the relative risk (**Fig. 8**) as a function of temperature (being the main driver of mortality). From this outcome, it is then possible to extract the specific minimum mortality temperature (T_{mm}), which may differ from place to place also reflecting different population

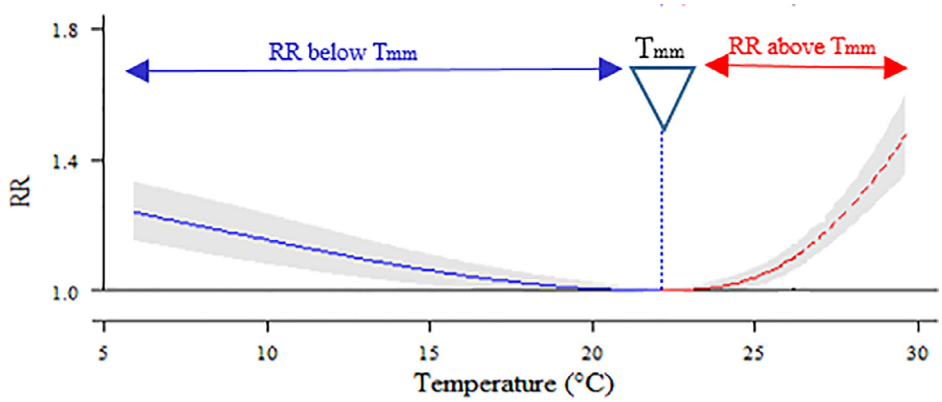

Fig. 8. Integrated relative risk curve for the city of Rome resulting from the D-MERF model.
Notes: *The graph represents the cumulative mortality risk curve of the D-MERF considering all exposures simultaneously - AT/T, O_3MA8 and PM_{10} - and accounting for time patterns, delayed effects, and non-linearities. Results are shown as a function of the temperature range.*

adaptability [29]. By summing up the daily adverse events associated with temperature above or below T_{mm}, it is then possible to obtain the total attributable mortality burden for the selected stressors (temperature and air pollutants). Finally, considering the T_{mm} as the optimal climatic point for human health, higher or lower temperatures in respect of this optimum value, can be referred to as “heat” and “cold” conditions, respectively.

The final RR-curve shows therefore the relative risk of mortality along the whole temperature distribution for Rome, which averages the information on the exposure-response association across the period 2004-2015. The graph shows higher relative risk coefficients for warm temperatures rather than cold ones.

The integrated function presented here could have been further complicated by inserting additional confounding factors or by considering not only temperature but also air pollution variables in a non-linear relationship with mortality. Indeed, several articles have recently attempted to model the relationship in question using nonlinear assumptions (e.g., [6,20]). However, the traditional linear association remains a plausible assumption and the most widespread (e.g., [18]). In addition, the use of non-linear health assessment models can make the interpretation of the exposure changes ‘more complex and a priori less predictable’ [20]. Especially in the case of multiple risk factors included in statistical-epidemiological models, it is always a good rule to weigh the marginal benefit obtained from the use of more complex models, in relation to the objective of the analysis, the input data, and the goodness of fit parameters. In the light of what said, and provided the lack of clear evidence on the non-linear relationship between air pollution and casualties in the historical observations for Rome, we believe our approach could provide a valid method to derive temperature induced adverse health effects when air pollutants are simultaneously at play, worsening the impact risk picture.

The procedure described here can offer a useful tool to investigate site-specific effects of climate on human health, accounting for air pollutant stressors. The relative risk curves and values must be considered a representation of the population responsiveness to the analyzed stressors. This method, applicable to other cities, may contribute to identifying priority areas at risk of adverse health effects determined by climate and/or air pollution conditions, in the context of global warming.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We are very grateful to Professor Antonio Gasparrini for his suggestions

Supplementary materials

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.mex.2022.101717.

References

[1] Airbase: Air quality e-reporting, https://www.eea.europa.eu/data-and-maps/data/aqereporting-8, last access: 15 May 2020.
[2] A. Analitis, K. Katsouyanni, A. Biggeri, M. Baccini, B. Forsberg, et al., Effects of cold weather on mortality: results from 15 European cities within the PHEWE project, Am. J. Epidemiol. 168 (2008) 1397–1408, doi:10.1093/aje/kwn266.
[3] A. Analitis, F. de'Donato, M. Scortichini, T. Lanki, X. Basagana, et al., Synergistic effects of ambient temperature and air pollution on health in Europe: results from the PHASE Project, Int. J. Environ. Res. Public Health. 15 (2018) 1856, doi:10.3390/ijerph15091856.
[4] B. Armstrong, Models for the relationship between ambient temperature and daily mortality, Epidemiology 17 (6) (2006) 624–631 November 2006, doi:10.1097/01.ede.0000239732.50999.8f.
[5] B. Armstrong, F. Sera, A.M. Vicedo-Cabrera, R. Abrutzky, D. OudinÅström, et al., The role of humidity in associations of high temperature with mortality: a multicountry, multicity study, Environ. Health Perspect. 127 (9) (2019), doi:10.1289/EHP5430.
[6] S. Bae, Y.H. Lim, S. Kashima, et al., Non-linear concentration-response relationships between ambient ozone and daily mortality, PLoS ONE 10 (6) (2015) e0129423 Published 2015 Jun 15, doi:10.1371/journal.pone.0129423.
[7] K. Bhaskaran, A. Gasparrini, S. Hajat, L. Smeeth, B. Armstrong, Time series regression studies in environmental epidemiology, Int. J. Epidemiol. 42 (2013) 1187–1195 2013, doi:10.1093/ije/dyt092.
[8] J.R. Buzan, K. Oleson, M. Huber, Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5, Geosci. Model Dev. 8 (2) (2015) 151–170, doi:10.5194/gmd-8-151-2015.
[9] J. Chen, G. Hoek, Long-term exposure to PM and all-cause and cause-specific mortality: Asystematic review and meta-analysis, Environ. Int. 143 (2020) 105974 (2020).
[10] F.K. de' Donato, M. Leone, M. Scortichini, M. De Sario, K. Katsouyanni, et al., Changes in the effect of heat on mortality in the last 20 years in nine European cities. results from the PHASE project, Int. J. Environ. Res. Public Health 12 (2015) 15567–15583 Int. J. Environ. Res. Public Health, doi:10.3390/ijerph121215006.
[11] D. Derwent, B. Forsberg, M. Amann, Health Risks of Ozone from Long-range Transboundary Air Pollution 2008, WHO, 2008 eBook ISBN: 9289042907 Print ISBN: 9289042893http://www.euro.who.int/__data/assets/pdf_file/0005/78647/E91843.pdf.
[12] J. Diaz, R. Carmona, I.J. Mirón, M.Y. Luna, C. Linaresa, Time trend in the impact of heat waves on daily mortality in Spain for a period of over thirty years (1983–2013), Environ. Int. 116 (2018) 10–17.
[13] A. Gasparrini, B. Armstrong, M.G. Kenward, Distributed lag non-linear models, Stat. Med. 29 (2010) 2224–2234 2010, doi:10.1002/sim.3940.
[14] A. Gasparrini, Distributed lag linear and non-linear models in R: the package dlnm, J. Stat. Softw. 43 (8) (2011) 1–20 2011.
[15] A. Gasparrini, Modelling lagged associations in environmental time series data: a simulation study, Epidemiology (Cambridge, Mass) 27 (6) (2016) 835–842, doi:10.1097/EDE.0000000000000533.
[16] H. Hersbach, B. Bell, P. Berrisford, et al., The ERA5 global reanalysis, QJR Meteorol. Soc. 146 (2020) 1999–2049 2020, doi:10.1002/qj.3803.
[17] C.Y. Guo, et al., Extensions of the distributed lag non-linear model (DLNM) to account for cumulative mortality, Environ. Sci. Pollut. Res. Int. 28 (29) (2021) 38679–38688, doi:10.1007/s11356-021-13124-0.
[18] J. Li, A. Woodward, X.Y. Hou, et al., Modification of the effects of air pollutants on mortality by temperature: a systematic review and meta-analysis, Sci. Total Environ. 575 (2017) 1556–1570 2017, doi:10.1016/j.scitotenv.2016.10.070.
[19] C. Liu, et al., Ambient particulate air pollution and daily mortality in 652 cities, N. Engl. J. Med. 381 (8) (2019) 705 2019.
[20] M.M. Nasari, M. Szyszkowicz, H. Chen, D. Crouse, M.C. Turner, M. Jerrett, R.T. Burnett, A class of non-linear exposure-response models suitable for health impact assessment applicable to large cohort studies of ambient air pollution, Air Qual. Atmos. Health 9 (8) (2016) 961–972, doi:10.1007/s11869-016-0398-z.
[21] I. Romieu, N. Gouveia, L.A. Cifuentes, A.P. de Leon, W. Junger, J. Vera, V. Strappa, M. Hurtado-Díaz, V. Miranda-Soberanis, L. Rojas-Bracho, L. Carbajal-Arroyo, G. Tzintzun-Cervantes, HEI Health Review Committee, Multicity study of air pollution and mortality in Latin America (the ESCALA study), Res. Rep. Health Eff. Inst. (171) (2012) 5–86.
[22] E. Samoli, A. Zanobetti, J. Schwartz, et al., The temporal pattern of mortality responses to ambient ozone in the APHEA project, J. Epidemiol. Community Health 63 (2009) 960–966 2009, doi:10.1136/jech.2008.084012.
[23] M. Scortichini, M. De Sario, F.K. de'Donato, M. Davoli, P. Michelozzi, M. Stafoggia, Short-term effects of heat on mortality and effect modification by air pollution in 25 Italian cities, Int. J. Environ. Res. Public Health 15 (2018) 1771, doi:10.3390/ijerph15081771.
[24] S.S. Stanišić, N. Stanišić, A. Stojić, Temperature-related mortality estimates after accounting for the cumulative effects of air pollution in an urban area, Environ. Health 15 (2016) 73, doi:10.1186/s12940-016-0164-6.
[25] R. Smith, Invited commentary: timescale-dependent mortality effects of air pollution, Am. J. Epidemiol. 157 (2003) 1066–1070 (2003).
[26] R.G. Steadman, Norms of apparent temperature in Australia, Aust. Meteorol. Mag. 43 (1994) 1–16 http://www.bom.gov.au/amm/papers.php?year=1994.
[27] A.M. Vicedo-Cabrera, F. Sera, C. Liu, B. Armstrong, A. Milojevic, et al., Short-term association between ozone and mortality: global two stage time series study in 406 locations in 20 countries, BMJ 368 (2020) m108 2020, doi:10.1136/bmj.m108.

[28] World Health Organization (WHO)Who Global Air Quality Guidelines: Particulate Matter ($Pm_{2.5}$ And Pm_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide, World Health Organization, 2021 https://apps.who.int/iris/handle/10665/345329 License: CC BY-NC-SA 3.0 IGO.
[29] Q. Yin, J. Wang, Z. Ren, J. Li, Y. Guo, Mapping the increased minimum mortality temperatures in the context of global climate change, Nat. Commun. 10 (1) (2019) 4640, doi:10.1038/s41467-019-12663-y.
[30] P. Yin, R. Chen, L. Wang, et al., Ambient ozone pollution and daily mortality: a nationwide study in 272 Chinese cities, Environ. Health Perspect. 125 (11) (2017) 117006 2017 NovPMID: 29212061; PMCID: PMC5947936, doi:10.1289/ehp1849.
[31] A. Zanobetti, J. Schwartz, Is there adaptation in the ozone mortality relationship: a multi-city case-crossover analysis, Environ. Health 7 (2008) 22, doi:10.1186/1476-069X-7-22.